

# Exhibit B

# Crystal Structures of an Antibody to a Peptide and Its Complex with Peptide Antigen at 2.8 Å

ROBYN L. STANFIELD, TERRY M. FIESER, RICHARD A. LERNER, IAN A. WILSON\*

The three-dimensional structures of an antibody to a peptide and its complex with the peptide antigen have been determined at 2.8 Å resolution. The antigen is a synthetic 19-amino acid peptide homolog of the C helix of myohemerythrin (Mhr). The unliganded Fab' crystals are orthorhombic with two molecules per asymmetric unit, whereas the complex crystals are hexagonal with one molecule per asymmetric unit. The Fab' and the Fab'-peptide complex structures have been solved independently by molecular replacement methods and have crystallographic *R* factors of 0.197 and 0.215, respectively, with no water molecules included. The amino-terminal portion of the peptide sequence (NH<sub>2</sub>-Glu-Val-Val-Pro-His-Lys-Lys) is clearly interpretable in the electron density map of the Fab'-peptide complex and adopts a well-defined type II β-turn in the concave antigen binding pocket. This same peptide amino acid sequence in native Mhr is α-helical. The peptide conformation when bound to the Fab' is inconsistent with binding of the Fab' to native Mhr, and suggests that binding can only occur to conformationally altered forms of the native Mhr or to apo-Mhr. Immunological mapping previously identified this sequence as the peptide epitope, and its fine specificity correlates well with the structural analysis. The binding pocket includes a large percentage of hydrophobic residues. The buried surfaces of the peptide and the antibody are complementary in shape and cover 460 Å<sup>2</sup> and 540 Å<sup>2</sup>, respectively. These two structures now enable a comparison of a specific monoclonal Fab' both in its free and antigen complexed state. While no major changes in the antibody were observed when peptide was bound, there were some small but significant side chain and main chain rearrangements.

THE RECENT STRUCTURE DETERMINATIONS OF FIVE MONOCLONAL antibody Fab fragments with their respective protein antigens (1-4) have revealed the detailed specificity of antibody-antigen interactions. In each case the antibodies have a high complementarity of fit with their antigens. While close examinations of these complexes have helped to answer many of the questions

concerning the structural basis of immune recognition, the question of whether antibody-antigen association can be described by a lock and key mechanism or by a handshake or induced fit mechanism is still the subject of debate. In order to answer such questions, the structures of both the free antibody and antigen must be determined in addition to the structure of the antibody-antigen complex. For the five reported complexes, the structures of the two antigens lysozyme and neuraminidase are known, but no structures of the respective free Fab's are available, although a partial structure of the unliganded D1.3-lysozyme Fab has been reported but not described (5).

Many monoclonal antibodies are not directed specifically against intact proteins. In particular, antibodies to synthetic antigens such as peptides constitute many of the monoclonal antibodies in use. Synthetic peptides corresponding in amino acid sequence to surface-accessible regions of a protein can elicit antibodies that recognize both peptide and protein (6). Antibodies to peptides have been used successfully to detect conformational changes in proteins (7) and for the purification of cloned and expressed proteins by fusing a short peptide epitope to their amino or carboxyl ends (8). Peptides have also been used to elicit neutralizing antibodies for foot-and-mouth disease (9), poliomyelitis (10), hepatitis B virus (11), malaria (12), and human immunodeficiency disease virus (13). While we are now beginning to understand the requirements for antibody-antigen recognition between antibodies to proteins and their antigens, almost nothing is known in structural terms about how an antibody to a peptide can bind both peptide and protein antigen.

Further understanding of antibody-peptide recognition should also allow us to improve on the design of synthetic peptides to induce catalytic antibodies capable of peptide bond hydrolysis. Successful catalytic antibodies to peptide analogues are known (14) as well as antibodies to various nonprotein immunogens including synthetic transition state analogues of organic reaction site interme-

**Table 1.** Data collection statistics for Fab'-peptide complex and native Fab' crystals. Data were collected with a Siemens-Nicolet-Xentronics area detector mounted on an Elliott GX-18 rotating anode x-ray generator operating at 40 kV, 55 mA, with 100-μm focusing optics. The Fab'-peptide complex data were collected from one crystal; the native Fab' data were collected from two crystals. Diffraction falls off rapidly at higher resolution as reflected in the *R<sub>sym</sub>* value.

Crystal form	Space group	Resolution (Å)	Observations (No.)	Unique reflections		Reflections >2.0 σ (No.)	<i>R<sub>sym</sub></i> (intensities)
				Measured (No.)	Expected (No.)		
Complex	P6 <sub>3</sub> 22	2.8	77906	26851	28967	22966	0.142
Free	P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	2.77	222264	25069	31299	24033	0.129

The authors are with the Department of Molecular Biology, Research Institute of Scripps Clinic, 10666 North Torrey Pines Road, La Jolla, California 92037.

\*To whom correspondence should be addressed.

**Table 2.** Rotational and translational parameters relating the light and heavy chains of the variable and constant domains. The elbow angle is the angle between the two pseudo-two-fold axes of the variable and constant domain. Domains were superimposed with the use of the program OVLAP (47). Elbow angles were calculated by the program ELBOW (48). The same set of 34 variable domain C $\alpha$  coordinates (4-6, 19-24, 34-39, 46-49, 68-70, 77-81, 90-93, 107-109 for the heavy chain and 4-6, 20-25, 33-38, 45-48, 63-65, 70-74, 86-89, 102-104 for the light chain) and 29 constant domain C $\alpha$  coordinates (120-123, 139-144, 153, 156-157, 171-173, 187-192, 208-211, 220-222 for the heavy chain and 114-117, 131-136, 145, 147-148, 159-161, 175-180, 194-197, 206-208 for the light chain) were used to superimpose the heavy chain onto the light chain for each of the three Fab' structures in the table. Coordinates used for the overlap are from conserved  $\beta$  sheet regions. The relative geometry of the light and heavy chains can be calculated in more than one way. For pairwise comparisons the entire V<sub>L</sub> domain (residues 3 to 110) of one Fab is superimposed onto the V<sub>L</sub> domain of the other structure, and the rotation and translation needed to

then superimpose the V<sub>H</sub> domains (3 to 112) is calculated. Such calculations (49) show that there are small but possibly real changes in the relative geometry of the unliganded and liganded Fab's. Relative rotations of 2.2° were required to refit the V<sub>H</sub> domains of either of the unliganded Fab' fragments onto the peptide bound Fab'. The comparable rotation relating the two unliganded Fab' fragments was 0.8°.

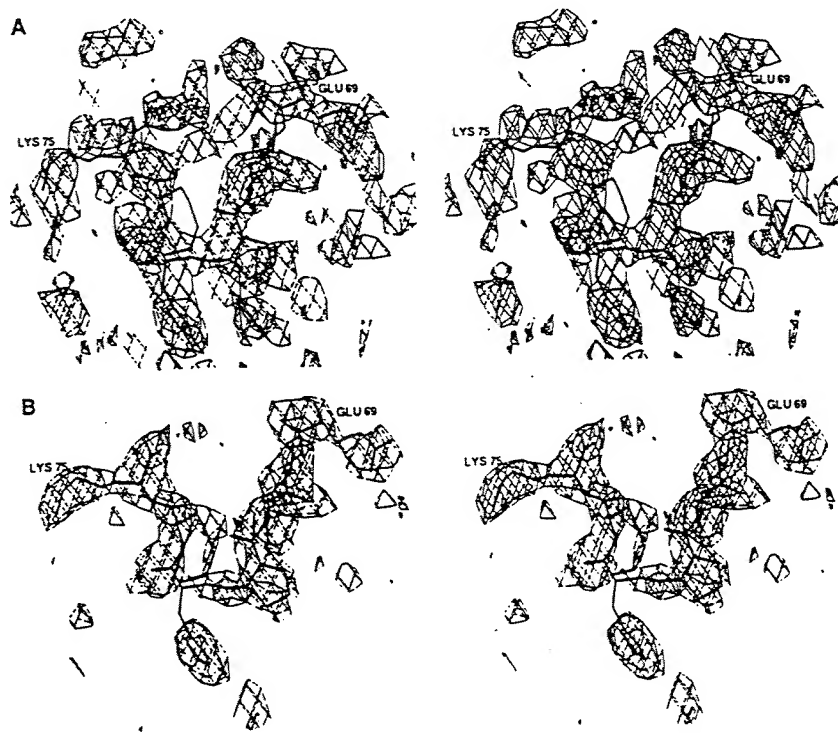
Fab'	Variable heavy to variable light		Constant heavy to constant light		Elbow angle (°)
	Rotation (°)	Translation (Å)	Rotation (°)	Translation (Å)	
B1312-peptide	173.5	0.44	167.2	1.84	157.1
B1312-free mol. 1	172.6	0.33	169.1	1.51	154.7
B1312-free mol. 2	174.2	0.57	169.4	1.69	156.4

diates (15). The variety of biological catalysts that can be produced by taking advantage of the immune repertoire as an almost infinite source of highly specific and highly complementary binding sites is therefore essentially limitless.

The importance of specific recognition of peptides by the immune system has been emphasized through the finding that foreign antigens are processed into short peptides before being presented by class I or class II major histocompatibility complex (MHC) antigens to the cellular T cell system. Crystal structures have been determined for two histocompatibility glycoproteins, HLA-A2 (16) and HLA-Aw68 (17). In each of these structures, electron density is seen in the antigen-binding cleft that presumably is that of a peptide antigen (16, 17). Based on sequence homology, proposals have been made that T cells share the immunoglobulin fold and recognize antigens in a manner similar to that of antibodies (18). Dual recognition of MHC and peptide by T cells may require their equivalent complementarity determining regions to be used for both interactions. Models indicate that residues corresponding to hypervariable heavy chain H3 and light chain L3 loops in T cells may play important roles in the peptide recognition event (18).

To better understand these mechanisms of antibody-peptide recognition we have systematically undertaken the structure determination of several antibodies to peptides and their peptide complexes (19, 20). We now describe one of a series of antibodies that recognize a synthetic homolog of the C helix of myohemerythrin (Mhr)—residues 69 to 87; EVVPHKKMKDFLEKIGGL (21)—and of apo-Mhr.

Mhr is a protein made up of four  $\alpha$  helices, whose structure has been solved and refined to high resolution (22). It has been the focus of several studies designed to predict areas of a protein that might be immunogenic (23, 24). Monoclonal antibodies to the synthetic Mhr C helix have been examined for affinity to peptide, apo-Mhr, and native Mhr (24). The epitopes on the peptide for each antibody were also determined. Three epitopes were localized; amino acid (aa) residues 69 to 73 (EVVPH), residues 79 to 84 (DFLEKI), and residues 86 to 87 (GL). The structure of this 19-aa synthetic peptide has been examined in solution by both circular dichroism (CD) and two-dimensional (2-D) nuclear magnetic resonance (NMR) spectroscopy (25). The NH<sub>2</sub>-terminal region of the peptide shows no stable secondary structure in water solution while the COOH-



**Fig. 1.** (A) Electron density corresponding to the bound Mhr C-helix peptide in a difference map calculated at 2.8 Å resolution with ( $F_o - F_c$ ) as coefficients and where the  $F_c$ 's were calculated with the use of only the coordinates for the Fab' B1312. The  $F_c$ 's were scaled to  $F_o$ 's in equivalent-sized shells of reciprocal space. The electron density was interpretable for the peptide sequence EVVPH. (B) Electron density of the peptide after subsequent refinement of the structure and building into the peptide omit map shown here provided a clear interpretation for the Lys<sup>74</sup> and Lys<sup>75</sup> side chains of the peptide. [Electron density is displayed with the use of the program FRODO (35) on a Silicon Graphics IRIS personal display system].

terminal region (residues 78 to 85) exists as a "nascent helix" (25).

We now describe the structures of an antibody to a peptide and its complex with its peptide antigen. The antigen, although relatively small compared to most proteins, forms many specific interactions with the Fab' which are not substantially different from those of the previously reported protein-Fab complexes (1-4). Conclusions drawn from the comparison of the free and peptide-bound Fab' structures reported here can now be added to our growing knowledge of the structural basis of antibody-antigen recognition.

**Crystallization and data collection.** Crystals were grown as described (19). The Fab'-peptide complex crystals grow in space group  $P6_322$  ( $a = b = 142.5$  Å,  $c = 101.5$  Å) with one Fab'-peptide complex per asymmetric unit. These crystals are grown from 1.85M phosphate at pH 5.75 and exhibit external hexagonal morphology. The Matthews coefficient ( $V_M$ ) of the crystals (26) is 2.99 Å<sup>3</sup> per dalton with the protein occupying 41 percent of the volume of the crystal, if we assume a protein partial specific volume of 0.74 cm<sup>3</sup>/g and 49,800 daltons as the molecular mass for the Fab'-peptide complex. The unliganded Fab' crystals grow in space group  $P2_12_12_1$  ( $a = 98.0$  Å,  $b = 151.7$  Å,  $c = 80.8$  Å) with two Fab' molecules per asymmetric unit. These rectangular rod-shaped crystals are grown from 1.1M sodium citrate, 1 percent methylpentanediol at pH 6.0. The native crystals have a  $V_M$  value of 3.16 Å<sup>3</sup> per dalton and the 49,760-dalton Fab' occupies 39 percent of the volume of the crystal. Both crystal forms diffract to at least 2.6 Å resolution although diffraction below 3.0 Å is weak. Nevertheless, both Fab' and Fab'-peptide complex crystals have long lifetimes in the x-ray beam (1 to 2 weeks on an Elliott GX-18 rotating anode operating at 40kV, 55 mA) and thus allow complete data sets to 2.8 Å resolution to be collected from just one crystal. X-ray diffraction data were collected (with a Siemens-Nicolet-Xentronics multi-wire area detector) and reduced [XENGEN package of programs (27) (Table 1)].

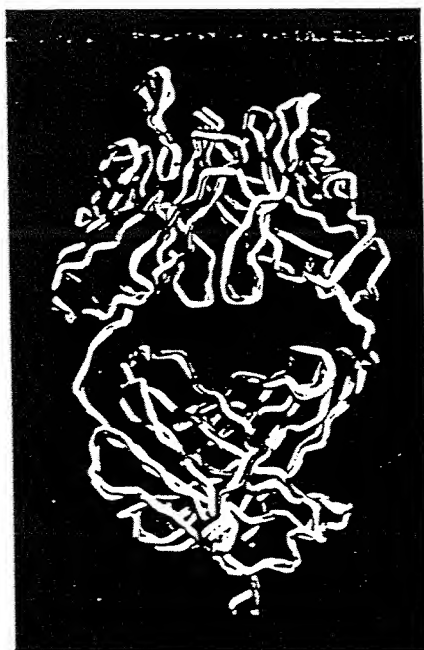


Fig. 2. Comparison of the free and peptide-complexed B1312 Fab' structures. The  $\alpha$  coordinates for the two Fab' molecules in the native asymmetric unit were superimposed onto the  $\alpha$  coordinates for the Fab' from the Fab'-peptide complex structure. The Fab' from the peptide complex structure is shown in white and the two unliganded Fab's in blue and yellow. The variable and constant domains were superimposed separately with the use of the program OVRLAP (47) (tube models calculated with the program MCS written by M. L. Connolly).

Table 3. Hydrogen bonds and salt bridge interactions in the Fab'-peptide complex. Those residues which have atoms close enough ( $<3.4$  Å) to form potential hydrogen bonds are listed. Seven of these hydrogen bonds are from the heavy chain and two from the light chain. The salt bridge involves two residues from the light chain. H1, H2, H3, L1, and L3 refer to heavy chain CDR's 1,2,3 and light chain CDR's 1, 3.

Peptide	Fab'	Distance (Å)
Glu <sup>69</sup> Oε1	Ser <sup>52</sup> Oγ (H2)	2.72
Glu <sup>69</sup> Oε1	Gly <sup>53</sup> N (H2)	3.38
Glu <sup>69</sup> Oε2	Ser <sup>55</sup> Oγ (H2)	2.85
Val <sup>70</sup> N	Ser <sup>52A</sup> Oγ (H2)	2.82
Val <sup>70</sup> O	Ser <sup>52A</sup> N (H2)	2.87
His <sup>73</sup> N	Pro <sup>99</sup> O (H3)	3.05
His <sup>73</sup> Nδ1	Gly <sup>91</sup> O (L3)	3.16
His <sup>73</sup> Nε2	Tyr <sup>95</sup> Oη (H3)	2.83
Lys <sup>75</sup> Nζ	Tyr <sup>32</sup> Oη (L1)	3.37
Lys <sup>75</sup> Nζ	Asp <sup>28</sup> Oδ1 (L1)	3.09
Lys <sup>75</sup> Nζ	Asp <sup>30</sup> Oδ1 (L1)	3.08

**Structure solution.** Both structures were solved by molecular replacement (MR). The Fab'-peptide complex structure was solved first and the Fab' portion of the structure was then used as a model to solve the free Fab' structure. For the Fab'-peptide complex structure solution, seven different Fab and Fab' fragments were tested as potential MR models. Coordinates for these models [NEWM (28), McPC603 (29), KOL (30), J539 (31), HyHEL-5 (3), and HyHEL-10 (4)] were taken from the Brookhaven Protein Data Bank. One additional coordinate set was obtained from within our laboratory [17/9 (32)].

Antibody Fab fragments consist of two structural domains connected by a flexible hinge or elbow. This elbow angle has been seen to vary from around 130° to 180° in known Fab structures (33). Because of this flexibility between the two domains, it is necessary for the rotation and translation functions to be calculated separately for each. The correct solutions are often still difficult to detect because of the nonideality of the model used in the MR calculations. The structure solution for the hexagonal crystal form was made especially difficult because each probe domain represented only 1/24th of the protein scattering matter in the crystal unit cell. Prior to rotation function calculations, the variable and constant domain coordinates for the seven Fab models were superimposed so that the planes containing the elbow angles were perpendicular to the z axis. Rotation functions were calculated for the different models with the use of the MERLOT programs for MR (34). For each domain, one rotation solution was found in common among the seven models. The variable domain of Fab KOL and the constant domain of Fab 17/9 were then chosen for the translation search. The translation

Table 4. van der Waals contact residues in the Fab'-peptide complex. The cutoff distance for van der Waals contacts was 4.1 Å. The heavy chain CDR's 1,2,3 and light chain CDR's 1,2,3 are represented by H1, H2, H3, L1, L2, L3. Six of the contact residues are from the light chain and 12 are from the heavy chain.

Peptide residues	Fab' contact residues
Glu <sup>69</sup>	Ser <sup>52</sup> (H2), Ser <sup>52A</sup> (H2), Gly <sup>53</sup> (H2), Ser <sup>55</sup> (H2)
Val <sup>70</sup>	Arg <sup>31</sup> (H1), Ser <sup>52</sup> (H2), Ser <sup>52A</sup> (H2), Pro <sup>99</sup> (H3)
Val <sup>71</sup>	Tyr <sup>56</sup> (H2), Phe <sup>58</sup> (H2)
Pro <sup>72</sup>	Ala <sup>33</sup> (H1), Ile <sup>51</sup> (H2), Ser <sup>52</sup> (H2), Phe <sup>58</sup> (H2), Tyr <sup>95</sup> (H3)
His <sup>73</sup>	Gly <sup>91</sup> (L3), Val <sup>94</sup> (L3), Pro <sup>96</sup> (L3), Tyr <sup>95</sup> (H3), Pro <sup>99</sup> (H3), Phe <sup>100</sup> (H3)
Lys <sup>74</sup>	Phe <sup>58</sup> (H2)
Lys <sup>75</sup>	Asp <sup>28</sup> (L1), Asp <sup>30</sup> (L1), Tyr <sup>52</sup> (L1), Phe <sup>100</sup> (H3)

problem was solved with the correlation coefficient search program BRUTE (35) in combination with packing information from the program PAKFUN in the MERLOT package (34). Models positioned with parameters derived from BRUTE were then subjected to several cycles of rigid body refinement with CORELS (36). During final cycles of rigid body refinement the variable heavy, variable light, constant heavy, and constant light chains were allowed to move as four separate rigid bodies in order to allow for differences from the input model in variable heavy to variable light or constant heavy to constant light pairings. The *R* factor after CORELS refinement for all data between 10.0 Å and 3.2 Å resolution was 0.451 percent.

This model was further refined with X-PLOR (37). Simulated annealing refinement with X-PLOR gave a model with an *R* factor of 0.26 for all data between 8.0 to 3.0 Å resolution. At this time, the nucleotide sequence for the variable domain of B13I2 became available (19). Constant domain sequences were taken from representative mouse IgG1 (immunoglobulin G-1) heavy chain and mouse kappa light chain sequences (38). The side chains of residues not conserved between the MR model and Fab' B13I2 were truncated to C $\beta$ 's, and the model was refined by simulated annealing for one cycle. Most of the truncated side chains were then clearly visible in a  $2F_o - F_c$  electron density map and rebuilt with the molecular model-building program FRODO (39).

Each hypervariable loop was then sequentially omitted from the

model during refinement and structure factor calculation and manually rebuilt. The entire model was then rebuilt into a series of "omit" maps where sequentially 10 percent of the structure was left out for a round of positional refinement and structure factor calculation. The *R* factor was then 0.24 for all data between 8.0 and 2.8 Å resolution. The last five residues of the light chain were not included in the model because they were not visible in the electron density map.

A difference electron density map with ( $F_o - F_c$ ) as coefficients now showed no significant features except for strong density (about  $5\sigma$  to  $6\sigma$ ) in the antigen binding site which was interpreted for nine residues of the 19-aa peptide antigen (Fig. 1A). The high quality of this electron density allowed us to immediately recognize the sequence EVVPH. This sequence is at the NH<sub>2</sub>-terminal end of the peptide (residues 69 to 73) and was previously considered from epitope mapping data to be the determinant of this antibody (24). The next four amino acids did not have easily identifiable side chain density and were built as alanines. These nine amino acids were incorporated into the model which was then refined with a round of simulated annealing refinement with X-PLOR. Group temperature factors where side-chain atoms and main-chain atoms of each residue have a different temperature factor were included at this point. The *R* factor after this refinement for data between 8.0 and 2.8 Å resolution greater than  $2\sigma$  was 0.215 (*R* = 0.23 for all data). At this point the side chain density for peptide residues 74 to 75 was interpretable but the side chains of residues 76 and 77 could not be unambiguously placed because of rather weak electron density. Current omit maps where the peptide has been deleted from a cycle of refinement and subsequent structure factor refinement have very clear electron density for the peptide through Lys<sup>75</sup> (Fig. 1B). In our current model only the sequence EVVPHKK is included. At present, the root-mean-square (rms) deviation from ideality for bond lengths is 0.022 Å and for bond angles is 4.40° for the Fab'-peptide complex.

The Fab' coordinates from the Fab'-peptide complex structure were then used to solve the native Fab' structure by MR. The volume of the unit cell of the unliganded Fab' crystal form indicated the presence of two Fab' molecules per asymmetric unit. A self-rotation function showed the presence of a noncrystallographic twofold symmetry axis. Cross-rotation functions between native

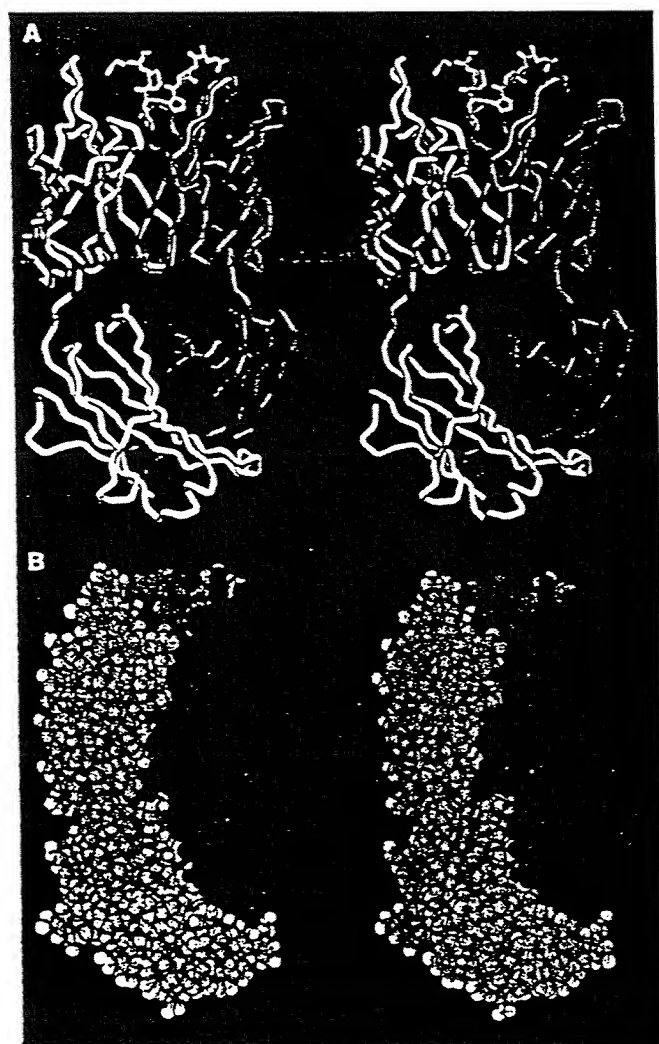


Fig. 3. (A) Stereoview of the structure of the Fab'-peptide complex. The dark blue chain corresponds to the heavy chain of the Fab' and the light blue chain corresponds to the light chain of the Fab'. Seven of the peptide residues of the Mhr C helix can be unambiguously placed in the electron density map at present and are shown in red bound in

the antigen binding pocket of the Fab'. The peptide has more interactions with the heavy chain than with the light chain. (B) Stereoview of the structure of the Fab'-peptide complex in a space-filling representation. The peptide (shown in red) can be seen embedded in the antigen binding pocket and sandwiched between Phe<sup>100</sup> and Phe<sup>58</sup> from the heavy chain. The Fab' is in the same orientation as shown in (A). Space-filling models were calculated with the program INSIGHT (Biosym Technologies on a Silicon Graphics Personal Iris). (C) Space-filling representation of the interaction of the peptide with the hypervariable loops (CDR's) of the Fab' as viewed looking into the antibody. Five of the six CDR's (H1, H2, H3, L1, and L3) interact with the peptide, whereas L2 does not make contact with the peptide. The peptide is shown in red and the CDR loops are colored as follows: L1, dark blue; L2, magenta; L3, green; H1, cyan; H2, pale magenta; H3, yellow.



data and models were calculated with MERLOT, and the two highest rotation solutions (one for each molecule in the asymmetric unit) were related by this noncrystallographic twofold axis. Crowther-Blow translation functions were calculated with MERLOT, and correct translations were verified by a BRUTE correlation coefficient search. The BRUTE model was then refined with CORELS with the variable light, variable heavy, constant light, and constant heavy chain from each Fab' in the asymmetric unit again being allowed to move as independent rigid bodies. The *R* factor after rigid body refinement was 0.355 for all data between 10.0 and 3.5 Å. The structure was then refined with X-PLOR to give an *R* factor of 0.203 for all data greater than 2 $\sigma$  between 10.0 and 2.8 Å after one round of simulated annealing refinement. Ten omit maps were calculated as previously described, and the model was rebuilt into these maps. The model was then refined with a round of

X-PLOR simulated annealing refinement to give an *R* factor after X-PLOR refinement with group temperature factors of 0.197 for all data greater than 2 $\sigma$  (*R* = 0.202 for all data). The five COOH-terminal residues of the light chain were visible in the electron density for each Fab' molecule and were included in the native model. At present, the rms deviations from ideality for bond lengths is 0.021 Å and for bond angles it is 4.22° for the free Fab'. The noncrystallographic twofold symmetry relating the two Fab' molecules in the native structure was not enforced during refinement. No water molecules have been included in either the native or complex structures. In both free Fab' and complex crystal forms additional density is also found attached to Asn<sup>26</sup> of the light chain, which is a potential glycosylation site (Asn<sup>26</sup>, Gln<sup>27</sup>, Thr<sup>27A</sup>).

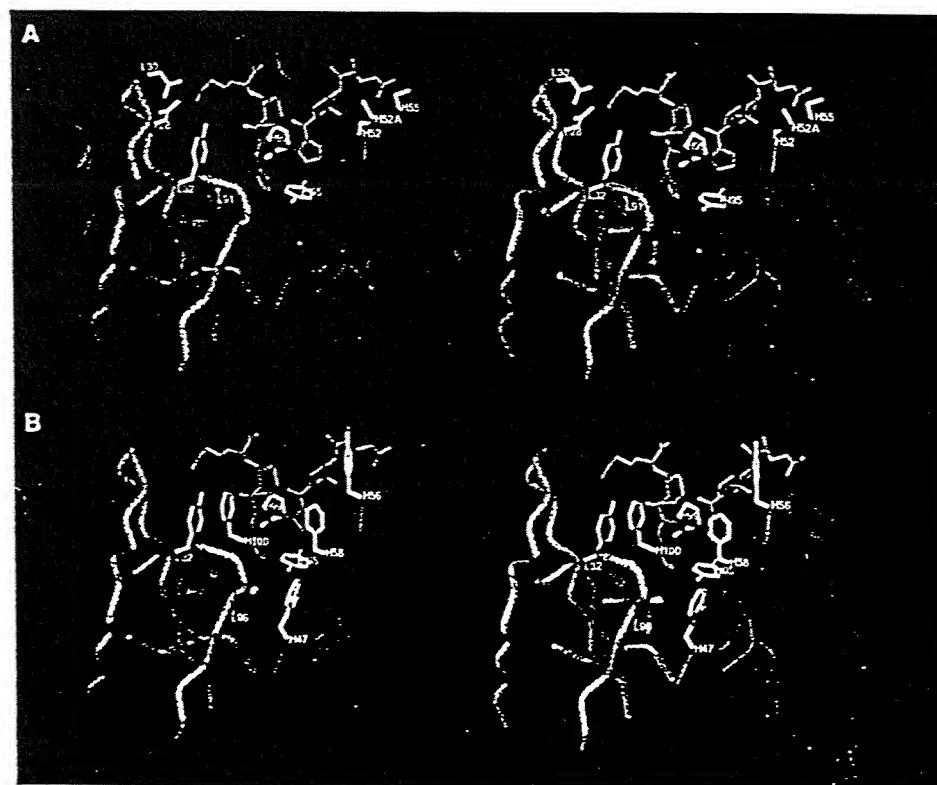
The solvent-accessible surface areas on the Fab' and the peptide were calculated with use of the program MS (40), with a 1.7 Å probe radius, and standard van der Waals radii (41). The cutoff for assignment of van der Waals contacts was 4.1 Å and the cutoff for hydrogen bonds was 3.4 Å. Numbering of the peptide-Fab' complex structure and the two independent Fab' molecules in the native structure is according to standard convention (38).

**Structure description.** The free and the peptide bound Fab' structures share the immunoglobulin fold found in all known Fab structures as judged by the polypeptide chain traces of the three molecules as shown superimposed in Fig. 2. The elbow angles for the Fab' are similar for the native and peptide bound forms and fall in the midrange of elbow angles seen previously in other Fab structures. The elbow angles and rotational and translational parameters relating the light and heavy chains of each domain for both crystal forms are shown in Table 2.

The peptide epitope consists of the first seven amino acids of the 19-chain which are bound in a concave pocket in the antibody combining site (Fig. 3). The peptide exists in a type II  $\beta$ -turn conformation involving residues Val<sup>71</sup>, Pro<sup>72</sup>, His<sup>73</sup>, and Lys<sup>74</sup>. The second and third residues of the turn are the most deeply buried

**Table 5.** The rms deviations (in angstroms) between the free and bound Fab' structures. The terms native 1 and native 2 refer to the two molecules in the asymmetric unit of the free (or unbound) structure. The light variable and the heavy variable domains were superimposed with Ca positions for Fab' framework residues as defined in Kabat *et al.* (38). The rms deviation is for only those framework Ca atoms. The light and heavy constant domains were superimposed with all Ca positions, and the rms deviations are for those atoms.

Comparison	Light		Heavy		
	Variable	Constant	Variable	Constant	
				All residues	Without residues 129-135
Native 1 vs. native 2	0.58	0.63	0.44	0.90	0.64
Native 1 vs. complex	0.47	0.70	0.53	1.35	0.89
Native 2 vs. complex	0.54	0.74	0.57	1.61	1.01



**Fig. 4.** (A) Stereoview of the Fab' residues which hydrogen bond to the peptide (red) in the antigen binding pocket. The side chains of Fab' residues that are involved in hydrogen bonding are shown in yellow. The Ca trace of the Fab' light chain is shown in light blue, the heavy chain in dark blue. (B) Stereoview of the hydrophobic and aromatic Fab' residues interacting with the peptide in the antigen binding pocket. The side chains of these Fab' residues are shown in yellow and the peptide in red. The light and heavy chains are colored as in (A). The residues in both (A) and (B) are numbered according to standard convention (38).

residues of the peptide. There are nine hydrogen bonds and a charge interaction between the Fab' and peptide. The charge interaction is between residue Lys<sup>75</sup> of the peptide and residues Asp<sup>28</sup> and Asp<sup>30</sup> from the light chain complementarity-determining region (CDR) L1 of the Fab'. Peptide residues Glu<sup>69</sup>, Val<sup>70</sup>, and His<sup>73</sup> are involved in hydrogen bonding to the Fab'. Residues Tyr<sup>32</sup> and Gly<sup>91</sup> from the light chain and Ser<sup>52</sup>, Ser<sup>52A</sup>, Gly<sup>53</sup>, Ser<sup>55</sup>, Tyr<sup>95</sup>, and Pro<sup>99</sup> from the heavy chain form hydrogen bonds with the peptide (Table 3).

Thus, residues directly involved in hydrogen bonding or charge-charge interactions with the peptide come from CDR's L1, L3, H2, and H3 (Fig. 4A). The seven peptide residues make 65 pairwise van der Waals contacts with the Fab' residues Asp<sup>28</sup>, Asp<sup>30</sup>, Tyr<sup>32</sup>, Gly<sup>91</sup>, Val<sup>94</sup>, and Pro<sup>96</sup> from the light chain and Arg<sup>31</sup>, Ala<sup>33</sup>, Ile<sup>51</sup>, Ser<sup>52</sup>, Ser<sup>52A</sup>, Gly<sup>53</sup>, Ser<sup>55</sup>, Tyr<sup>56</sup>, Phe<sup>58</sup>, Tyr<sup>95</sup>, Pro<sup>99</sup>, and Phe<sup>100</sup> from the heavy chain (Table 4). These residues are from CDR's L1, L3, H1, H2, and H3 as well as a few framework residues. Twelve of the van der Waals contacts are with the light chain, and 53 of the van der Waals contacts are with the heavy chain. There are several residues that are buried by the binding of the peptide but are not within van der Waals contact distance: Leu<sup>27D</sup> from the light and Cys<sup>32</sup>, Trp<sup>47</sup>, Gly<sup>40</sup>, Ser<sup>96</sup>, Ser<sup>97</sup>, and Asp<sup>98</sup> from the heavy chain. As in several other antibodies and a light chain dimer [HyHEL-5 (3), D1.3 (1), McPC603 (29), TE32, TE33 (42), and Mcg (43)] there is a high percentage of aromatic residues in the B1312 antigen binding pocket (Fig. 4B). The buried surface area on the Fab' is 540 Å<sup>2</sup> and the peptide buried surface covers 460 Å<sup>2</sup>, an indication that the surfaces are highly complementary although the fit is not exact (Fig. 5). Of the buried Fab' surface, 78 percent consists of heavy chain residues while only 22 percent of the surface is attributable to light chain residues.

The rms deviation between the two native variable light chains for 80 Cα atoms in the framework region (1 to 23, 35 to 49, 57 to 88, 98 to 107) is 0.58 Å. The corresponding rms deviation between the two native variable heavy chains for 85 Cα atoms in the framework region (1 to 30, 36 to 49, 66 to 94, 103 to 113) is 0.44 Å. The rms deviation between the two native constant light chains is 0.63 Å and between the constant heavy chains is 0.90 Å. The heavy chain of the constant regions appears to be disordered between residues 129 to 135; electron density for this region is poor, and the region has higher than average temperature factors. Deletion of residues 129 to 135 from the overlap and rms deviation calculation gives an rms deviation of 0.64 Å for the two native constant heavy chains. The rms deviations between each of the molecules in the native crystal form and the complexed Fab' are similar to the deviations seen between the two native Fab' fragments in the same asymmetric unit (Table 5).

**Antibody-antigen interactions.** The above structural data provide further information that should help resolve some of the fundamental questions of immune recognition. The question of whether a peptide immunogen adopts a defined secondary structure or any regular structure at all is addressed by analysis of the Fab'-peptide complex structure. Whether the conformation of the peptide in the Fab'-peptide complex is the same as the peptide conformation in the intact protein can now be answered for this particular immune complex. The previously unresolved question of whether an antibody changes conformation upon antigen binding is answered on the basis of the structure determinations of the free and complexed Fab' structures.

The antigen binding regions of the two free Fab' molecules in the same asymmetric unit and the peptide bound Fab' molecule are shown superimposed by domain in Fig. 2. There are no large main chain deviations among the three molecules and there appear to be only small side chain and main chain differences between the two free Fab' molecules as well as between the peptide-bound Fab' and

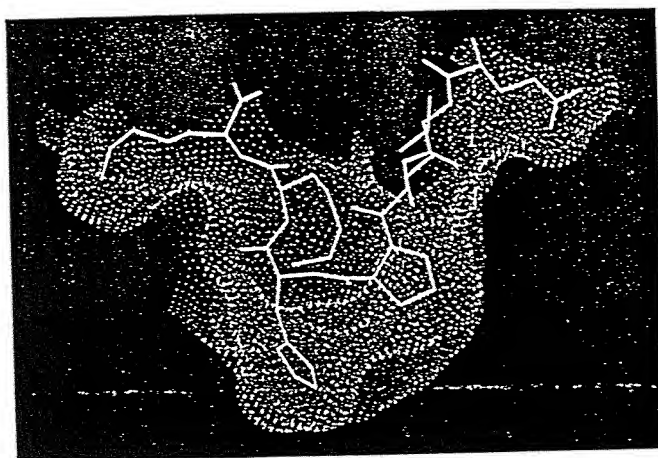


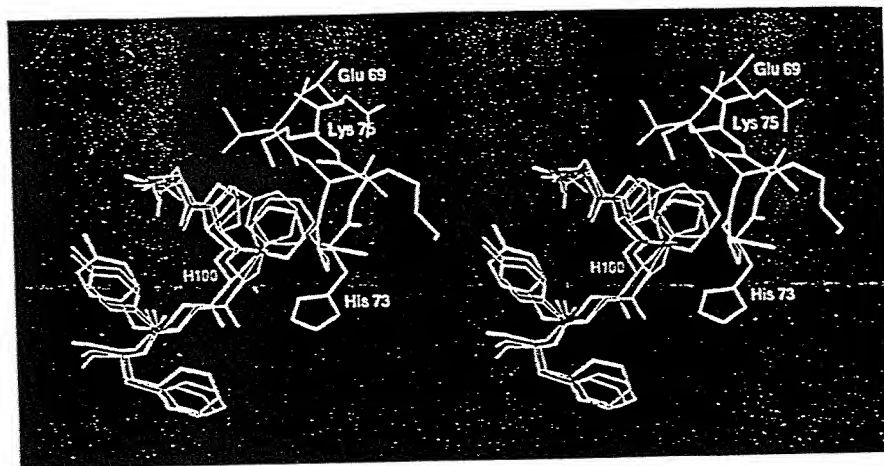
Fig. 5. Comparison of the peptide and Fab' surfaces buried in the antigen-antibody complex. The solvent accessible surfaces of the peptide and the Fab' that are buried upon peptide binding are complementary in size and shape. The area of the buried surface for peptide is 460 Å<sup>2</sup> and for Fab' is 540 Å<sup>2</sup>. Only the coordinates for the peptide are shown for clarity. Buried surfaces were calculated with the program MS (38) and displayed with the program FRODO (35) on a Silicon Graphics Personal IRIS.

the free Fab' molecules. The most striking side chain movement in the Fab'-peptide complex structure is that of Phe<sup>100</sup> from the heavy chain (Fig. 6). The center of gravity of the ring has moved 1.74 Å from the center of gravity of the ring in native molecule 1 and 2.13 Å from the center of gravity of the ring in native molecule 2 in a comparison of the superimposed coordinate sets. The Cα positions in this region of the heavy chain have also moved with respect to their equivalent positions in the native molecules; Cα for Pro<sup>99</sup> has moved an average of 1.24 Å, Cα for Phe<sup>100</sup> has moved an average of 1.20 Å, and the Cα for Tyr<sup>100B</sup> has moved an average of 1.04 Å. The movement of this section of main chain and the Phe<sup>100</sup> ring is away from the peptide binding site presumably to permit the peptide to bind. If the peptide is positioned in the structures of the native molecules in this superimposed orientation, there are close contacts between the peptide and this loop. There are also main chain variations of up to 2.0 Å in the CDR loop L1. However, the deviation between the two native L1 loops is as large as that seen when either native L1 loop is compared to the complex L1 loop. This long L1 loop appears somewhat disordered in both native and complex electron density maps and refines with higher than average temperature factors. The environment of this loop also differs in each of the free and complex structures. The corresponding loop in McPC603 also appears to have a disordered structure (29).

The areas of the buried surfaces of the Fab' and the peptide are 540 and 460 Å<sup>2</sup>, respectively, which are perhaps larger than was expected for Fab-peptide interactions. These areas are only slightly smaller than those seen previously for antibody-protein complexes [HyHEL-5, 750 Å<sup>2</sup> for both antibody and lysozyme; HyHEL-10, 720 Å<sup>2</sup> for the antibody and 774 Å<sup>2</sup> for lysozyme; and D1.3, 690 Å<sup>2</sup> for the antibody and 680 Å<sup>2</sup> for lysozyme (44)]. Corresponding buried surface areas for antibody-hapten complexes are, as expected, smaller [McPC603, 161 Å<sup>2</sup> for the antibody and 137 Å<sup>2</sup> for phosphocholine; and Fab 4-4-20, 308 Å<sup>2</sup> for the antibody and 266 Å<sup>2</sup> for fluorescein (44)].

Comparison of these free and bound Fab' B1312 structures further supports the theory that variations in the elbow angle are due merely to flexibility of the antibody and are not a result of antigen binding (44). The variation in elbow angles between two Fab' fragments in the same crystal asymmetric unit is comparable to the difference in elbow angle between the native Fab' and the

Fig. 6. Stereoview of a conformational difference between the free and peptide-bound Fab'. The H3 CDR around Phe<sup>100</sup> displays the largest differences between the unbound and the bound Fab' that can be attributed to the peptide binding. This region of the two native molecules is shown in green superimposed on the same region of the peptide-Fab' complex (red). The peptide is shown in yellow. Displacements of the backbone and side chain atoms are apparent. For example, the ring of Phe<sup>100</sup> has moved from its position in the native Fab' by more than 1 Å in the peptide-Fab' complex. Local movements in this H3 loop appear to be necessary to avoid a steric clash when the peptide binds to the native. Differences in the L1 loop can also be seen in Fig. 2. However, the L1 loops of the two unbound Fab' fragments differ from each other more than they do from the complex. Hence, no significance can be placed on deviations in this L1 loop conformation between the free and bound Fab' structures. The coordinates were displayed with the program FRODO (35) on a Silicon Graphics Personal IRIS.



peptide bound Fab' (Table 2). Small differences can also be calculated for the relations between the variable light and variable heavy domains when the unliganded Fab' fragments are compared with the peptide complexed Fab'. These differences in the relative geometrical disposition of the ligand-free versus ligand-bound Fab' fall in the range of 0° to 2° (Table 2). As these calculations are sensitive not only to the actual residues used in the superposition analyses but also to the comparison method, we cannot be certain at present of any variation in the variable heavy to variable light pairings due to binding of antigen. Local conformational differences, for example, in loop regions, can also affect the numerical results. Further examination of the variable heavy to variable light pairings will be made after further refinement with higher resolution data.

The buried surface of Fab' B1312 and that of the peptide are very complementary with only a few small pockets on the antibody surface not filled by the peptide. The calculated buried surfaces are shown in Fig. 5. The number of hydrogen bonds, salt bridges, and van der Waals contacts seen in the peptide-Fab' complex are similar to those seen in the HyHEL-5-lysozyme complex (3) and the HyHEL-10-lysozyme complex (4) (nine hydrogen bonds, one bifurcated salt link, and 65 van der Waals contacts for B1312 compared to ten hydrogen bonds, three salt links, and 74 van der Waals contacts for HyHEL-5; and 14 hydrogen bonds, one salt link, and 111 van der Waals contacts for HyHEL-10).

The residues of the bound peptide that interact with the Fab' as seen in the x-ray structure are in excellent agreement with epitope mapping results from an earlier immunological study (24). Replacement series ELISA's (enzyme-linked immunosorbent assays) where each residue of the epitope was replaced by each of the other 19 amino acids showed that Glu<sup>69</sup>, Pro<sup>72</sup>, and His<sup>73</sup> are essential for binding; peptides where these residues were replaced showed very low levels of binding to antibody. In fact, these residues are involved in specific hydrogen bond interactions in the peptide-Fab' complex (Table 3 and Fig. 4A). Replacement series ELISA's showed Val<sup>70</sup> and Val<sup>71</sup> to be less essential; other side chains can be substituted for these residues, especially Val<sup>70</sup>, while maintaining fairly high levels of binding. This is in agreement with the observation that the side chains of these two residues contribute relatively few specific interactions from the peptide to the antibody (Table 3 and Fig. 4). In the same study, significant binding was seen between antibody B1312 and either native Mhr or apo-Mhr in competition ELISA experiments. However, the apo-Mhr was found to bind a factor of 80 times more tightly than native Mhr. After examination of the

Fab'-peptide structure, we found it hard to imagine how this Fab' could bind native Mhr without substantial rearrangement of the determinant residues in Mhr. In fact, in the peptide-Fab' complex, peptide residue His<sup>73</sup> is intimately involved in the binding pocket whereas in native Mhr, His<sup>73</sup> is coordinated to an iron atom internally and would be unavailable for binding to antibody. In light of these structural results, as well as further experimentation (45), we now suggest that the positive competition (ELISA) results with native Mhr were brought about by conformational changes in the solution phase Mhr due to unknown causes. This observation is in agreement with the increased level of binding between B1312 and apo-Mhr. In order to answer these questions one must solve the structure of the antibody while bound to intact Mhr. Current efforts to co-crystallize the highly interesting Fab'-apo-Mhr complex seem promising.

The structure of the 19-aa long C-helix peptide has also been studied in solution by NMR and CD. We can now compare the structures of the peptide in solution, in the native protein and while bound to an antibody. In solution, the peptide shows no regular structure in the NH<sub>2</sub>-terminal region. The COOH-terminal region of the peptide exists as a nascent helix; the peptide converts rapidly between a series of turnlike structures (25). In the native protein, the entire peptide adopts an  $\alpha$ -helical conformation (22) with His<sup>73</sup> and His<sup>77</sup> coordinated to the iron atoms in Mhr. In the crystal structure, the first nine amino acids of the peptide can be seen with the remainder of the peptide not visible and presumed to be disordered. As the side chains of Met<sup>76</sup> and His<sup>77</sup> are not clear, these residues are not included in the model at present. This NH<sub>2</sub>-terminal region of the peptide forms a type II  $\beta$ -turn while bound to antibody and is significantly different from either the peptide solution structure or the peptide structure in the native protein. It is interesting that although the peptide conformation that is recognized by the antibody differs from that in the native protein, the antibody binds to the peptide in a well-defined secondary structure commonly found in proteins.

A 6-aa sequence of the heavy chain constant region (129 to 135) of both the peptide bound Fab' and the native Fab' was especially difficult to visualize in the electron density maps, presumably due to disorder of the region. Examination of the temperature factors of other antibody structures deposited in the Brookhaven Data Bank revealed a trend for larger than average temperature factors in this region. There is also higher than average sequence variability in this region according to the variability index for constant domain sequences as calculated by Kabat *et al.* (38).



More antibody-antigen complex structures are needed for comparison with the native antibody structures before we can map the full range of antibody-antigen interactions. These could include, for example, the structure of other monoclonal antibodies to peptides (19, 20) such as B13A2 (46), which recognizes the COOH-terminal epitope of the Mhr C-helix peptide. A comparison of the structure of the same C-helix peptide while bound to different Fab' fragments at two different epitopes should provide structural data to show the effect of environment on peptide conformation. Precisely how an antibody can bind to the same sequence in both a peptide and a protein remains to be determined.

#### REFERENCES AND NOTES

1. A. G. Armit, R. A. Mariuzza, S. E. V. Phillips, R. J. Poljak, *Science* 233, 747 (1986).
2. P. M. Colman et al., *Nature* 326, 358 (1987); P. M. Colman et al., *Philos. Trans. R. Soc. Lond. Ser. B* 323, 511 (1989).
3. S. Sheriff et al., *Proc. Natl. Acad. Sci. U.S.A.* 84, 8075 (1987).
4. E. A. Padlan et al., *ibid.* 86, 5938 (1989).
5. G. A. Bentley et al., *Cold Spring Harbor Symp. Quant. Biol.* 54, 239 (1990).
6. R. A. Lerner, *Adv. Immunol.* 36, 1 (1984).
7. J. M. White and I. A. Wilson, *J. Cell Biol.* 105, 2887 (1987).
8. J. Field et al., *Mol. Cell. Biol.* 8, 2159 (1988).
9. M. J. Francis et al., in *Vaccines '87*, R. M. Chanock, R. A. Lerner, F. Brown, H. Ginsberg, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1987), p. 60.
10. M. Chow, J. L. Bittle, J. Hogle, D. Baltimore, in *Modern Approaches to Vaccines*, R. M. Chanock and R. A. Lerner Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1984), p. 257.
11. D. R. Milich, A. McLachlan, J. L. Hughes, A. Moriarty, G. B. Thornton, in *Vaccines '88*, H. Ginsberg, F. Brown, R. A. Lerner, R. M. Chanock, Eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988), p. 13.
12. M. E. Patarroyo et al., *Nature* 332, 158 (1988).
13. K. Javaherian et al., *Proc. Natl. Acad. Sci. U.S.A.* 86, 6768 (1989).
14. B. L. Iverson and R. A. Lerner, *Science* 243, 1184 (1989).
15. R. A. Lerner and S. J. Benkovic, *BioEssays* 9, 107 (1988).
16. P. J. Bjorkman et al., *Nature* 329, 506 (1987); *ibid.*, p. 512.
17. T. P. J. Garrett, M. A. Saper, P. J. Bjorkman, J. L. Strominger, D. C. Wiley, *ibid.* 342, 692 (1989).
18. M. M. Davis and P. J. Bjorkman, *ibid.* 334, 395 (1989).
19. E. A. Stura et al., *J. Biol. Chem.* 264, 15721 (1989); J. M. Rini, R. L. Stanfield, E. A. Stura, U. Schulze-Gahmen, I. A. Wilson, in *Use of X-Ray Crystallography in the Design of Antiviral Agents*, W. G. Laver and G. M. Air, Eds. (Academic Press, San Diego, CA, 1990), p. 87.
20. U. Schulze-Gahmen et al., *J. Biol. Chem.* 263, 17100 (1988).
21. Abbreviations for the amino acid residues are: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
22. S. Sheriff, W. A. Hendrickson, J. L. Smith, *J. Mol. Biol.* 197, 273 (1987).
23. J. A. Tainer et al., *Nature* 312, 127 (1984).
24. T. M. Fieser, J. A. Tainer, H. M. Geysen, R. A. Houghten, R. A. Lerner, *Proc. Natl. Acad. Sci. U.S.A.* 84, 8568 (1987).
25. H. J. Dyson, M. Rance, R. A. Houghten, P. E. Wright, R. A. Lerner, *J. Mol. Biol.* 201, 201 (1988).
26. B. W. Matthews, *ibid.* 33, 491 (1968).
27. A. J. Howard et al., *J. Appl. Cryst.* 20, 383 (1987).
28. F. A. Saul, L. M. Amzel, R. J. Poljak, *J. Biol. Chem.* 253, 585 (1978).
29. Y. Satow et al., *J. Mol. Biol.* 190, 593 (1986).
30. M. Marquart, J. Deisenhofer, R. Huber, W. Palm, *ibid.* 141, 369 (1980).
31. S. W. Suh et al., *Protein Struct. Funct. Genet.* 1, 74 (1986).
32. J. M. Rini and I. A. Wilson, unpublished results.
33. S. Sheriff et al., in *Structure and Expression*, R. H. Sarma, and M. H. Sarma, Eds. (Adenine Press, Albany, NY, 1988), vol. 1, p. 49.
34. P. M. D. Fitzgerald, *J. Appl. Cryst.* 21, 273 (1988).
35. M. Fujinaga, and R. Read, *ibid.* 20, 517 (1987).
36. J. L. Sussman, S. R. Holbrook, G. M. Church, S.-H. Kim, *Acta Cryst.* A33, 800 (1977).
37. A. T. Brünger, J. Kuriyan, M. Karplus, *Science* 235, 458 (1987).
38. E. A. Kabat, T. T. Wu, M. Reid-Miller, H. M. Perry, K. S. Gottesman, *Sequences of Proteins of Immunological Interest* (National Institutes of Health, Bethesda, MD, ed. 4, 1987).
39. A. T. Jones, *J. Appl. Cryst.* 11, 268 (1978).
40. M. L. Connolly, *ibid.* 16, 548 (1983).
41. D. A. Case and M. Karplus, *J. Mol. Biol.* 132, 343 (1979).
42. R. Levy, O. Assulin, T. Scherf, M. Levitt, J. Angilister, *Biochemistry* 28, 7168 (1989).
43. M. Schiffer, R. L. Girtling, K. R. Ely, A. B. Edmundson, *ibid.* 12, 4620 (1973).
44. D. R. Davies, E. A. Padlan, S. Sheriff, *Annu. Rev. Biochem.* 59, 439 (1990).
45. B. D. Spangler, unpublished results; I. A. Wilson et al., unpublished results.
46. D. H. Fremont and I. A. Wilson, unpublished results.
47. M. G. Rossmann and P. Argos, *J. Biol. Chem.* 250, 7525 (1975).
48. D. H. Fremont and T. O. Yeates, unpublished results.
49. A. M. Lesk, in *Biosequences: Perspectives and User Services in Europe*, C. Saccone Ed. (European Economic Commission, Strasbourg, 1986), pp. 23-28.
50. We thank R. Samodal for technical assistance in the Fab' and Mhr preparations and binding assays; Drs. B. Spangler and F. Stevens of Argonne National Laboratory for helpful discussions and unpublished binding data; R. Ghadiri for help with spectroscopic studies; R. Houghten for peptide synthesis; E. Stura for crystallizing the Fab'-peptide complex crystals and other crystallization assistance; and J. Rini, J. Arevalo, D. Filman, J. Hogle, D. Fremont, A. Lesk, and C. Chothia for discussions. Supported by NIH grant GM38794 (I.A.W.) and NIH training fellowship AI-07244 (R.L.S.). This is publication #6173-M3 from The Research Institute of Scripps Clinic.

27 December 1989; accepted 9 April 1990

## AAAS-Newcomb Cleveland Prize

### To Be Awarded for an Article or a Report Published in *Science*

The AAAS-Newcomb Cleveland Prize is awarded to the author of an outstanding paper published in *Science*. The value of the prize is \$5000; the winner also receives a bronze medal. The current competition period began with the 2 June 1989 issue and ends with the issue of 25 May 1990.

Reports and Articles that include original research data, theories, or syntheses and are fundamental contributions to basic knowledge or technical achievements of far-reaching consequence are eligible for consideration of the prize. The paper must be a first-time publication of the author's own work. Reference to pertinent earlier work by the author may be included to give perspective.

Throughout the competition period, readers are invited to nominate papers appearing in the Reports or Articles sections. Nominations must be typed, and the following information provided: the title of the paper, issue in which it was published, author's name, and a brief statement of justification for nomination. Nominations should be submitted to the AAAS-Newcomb Cleveland Prize, AAAS, Room 924, 1333 H Street, NW, Washington, DC 20005, and must be received on or before 30 June 1990. Final selection will rest with a panel of distinguished scientists appointed by the editor of *Science*.

The award will be presented at the 1991 AAAS annual meeting. In cases of multiple authorship, the prize will be divided equally between or among the authors.